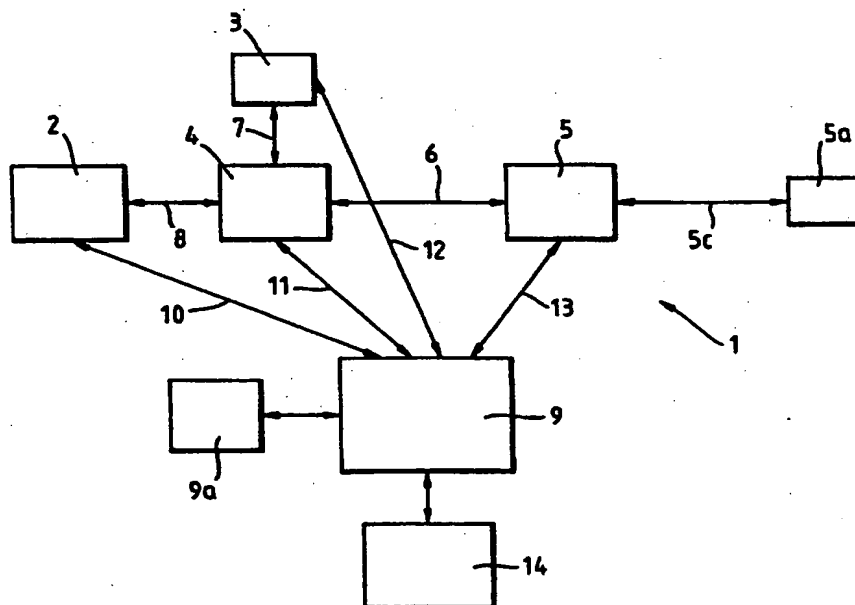# PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04L 12/56, H04Q 11/04	A1	(11) International Publication Number: WO 95/28787
		(43) International Publication Date: 26 October 1995 (26.10.95)

(21) International Application Number: PCT/GB95/00823

(22) International Filing Date: 11 April 1995 (11.04.95)

(30) Priority Data:
94302619.5   13 April 1994 (13.04.94)   EP
(34) Countries for which the regional or
    international application was filed: AT et al.

(60) Parent Application or Grant
    (63) Related by Continuation
    US   08/252,097 (CIP)
    Filed on   31 May 1994 (31.05.94)

(71) Applicant (for all designated States except US): BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): KEY, Peter, Bernard [GB/GB]; Bredfield Lodge, 30 Bredfield Road, Woodbridge, Suffolk IP12 1JB (GB). GRIFFITHS, Thomas, Rhodri [GB/GB]; 280 Alexandra Road, Ipswich, Suffolk IP4 2RL (GB). GIBBENS, Richard, John [GB/GB]; Lyndewode

Research Limited, 23 Waterhouse, Waterlane, Cambridge CB4 1XZ (GB). KELLY, Francis, Patrick [GB/GB]; Lyndewode Research Limited, 23 Lyndewode Road, Cambridge CB1 2HN (GB).

(74) Agent: MORGAN, Marc; BT Group Legal Services, Intellectual Property Dept., 13th floor, 151 Gower Street, London WC1E 6BA (GB).

(81) Designated States: AU, CA, CN, JP, KR, NZ, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

(54) Title: A COMMUNICATION NETWORK CONTROL METHOD AND APPARATUS

(57) Abstract

A communication network (1) comprises a number of network nodes (2, 3, 4 and 5) linked together by transmission paths (6, 7 and 8). A network manager (9) controls acceptance of calls onto transmission paths by classifying an incoming call and selecting from the database (14) an appropriate mean to peak bit rate distribution. This selected distribution is then used with the monitored transmission path load to determine a posterior distribution which reflects the load. The decision to accept or reject the call is made on the basis of the posterior distribution by a call acceptance algorithm considering the effective bandwidth for example.

A COMMUNICATION NETWORK CONTROL METHOD AND APPARATUS

This invention relates to a communication network control method and apparatus.

In order to maintain the quality of service in communication networks, it is necessary to limit the number of calls allowed onto the network. The optimum number of calls to be allowed on the network is determined by the characteristics of the network which depend on trunk speeds and other hardware constraints, such as buffer sizes. The mix of call types and the relative number of each is also an important consideration.

Where a user of a network knows the mean bit rate of the call he wishes to place on the network, it is possible to use methods which employ Convolution or Large Deviation Approximation techniques to determine whether the call can be carried by the network without the quality of service dropping below a minimum value. However, in practice the user is unlikely to know what the call's mean bit rate will be. This will be especially the case where the network carries broadband services. It has been proposed to monitor the network's traffic load and to refine the control method in a dynamic way. This type of method has been called dynamic Connection Admission Control (CAC).

Given the uncertainties surrounding the data type that will be carried on communication networks offering broadband services, there is a generally held belief in the communications field that it will not be possible to employ statistical multiplexing techniques on networks that employ asynchronous transfer mode techniques (ATM).

The present invention arose from a realisation by the inventor that whilst a resource trying to make a connection with the network would be unlikely to know its mean bit rate, it would undoubtedly have knowledge about its peak bit rate.

According to the invention there is provided a communication network control method for controlling acceptance of a call from a first network node to a second network node over a transmission path between the nodes comprising providing a database of activity distributions for particular classes of calls, classifying the call,

5   selecting from the database a distribution appropriate to the call, determining a current load on the transmission path due to calls of that class, deriving a posterior activity distribution from the selected distribution and current load, and rejecting or accepting the call on the basis of the posterior activity distribution.

10

By providing a database of activity distributions for particular classes of calls and generating, from this, and a measure of the current load due to calls of that class, a posterior activity distribution, a much better estimate of the activity can be obtained.

15

An activity distribution could be a variance distribution or various forms of bit rate distribution. It can be thought of as the likelihood or probability of a certain call load being attained by a particular class of call. Preferably, the activity distribution is a mean to peak bit rate distribution.

20

The posterior distribution will then be used by a call acceptance control algorithm to determine whether the call can be accepted onto the network. The preferred way of doing this is to determine the effective bandwidth of the current load on the transmission path from the posterior distribution. From the effective

25   bandwidth, the quality of service if the call is accepted will be determined and compared with a required quality of service. If the quality of service is maintained then the call is accepted, if not, the call is rejected for that transmission path.

Preferably, the posterior distribution is generated by a Bayesian analysis of the

30   selected distribution and the current load.

Preferably, the call is classified according to a parameter declared by the call. In a preferred embodiment the parameter is the peak bit rate of the call. Other

parameters that may be used are much as whether its a particular type of call for example video or using a particular type of equipment.

In some networks, certain classes of calls may be allocated certain bandwidth out
5 of the total transmission path bandwidth. However, in the other networks . there will not be any collective bandwidth allocation. Whilst in the former form of network it will be possible to compare the posterior distribution for the call class with the allocated bandwidth for that class and use this as the basis for whether or not the call is accepted, in the latter type of network, preferably, the load on the
10 transmission path due for all call classes is periodically monitored and a measure of the monitored load is stored. The call is then accepted on the basis of the stored monitored load and the posterior distribution by, for example determining the effective bandwidth of the stored load and posterior distribution and comparing this with the capacity of the transmission path to ensure that the required quality
15 of service is maintained.

The advantage of periodically monitoring the load on the transmission path over producing, for example, a posterior distribution for each call class, upon a request being received to allow a call of a particular class onto the transmission path, is
20 that of speed in reaching the decision whether to accept or reject the call.

According to a second aspect of the invention there is provided apparatus for controlling acceptance of a call from a first network node to a second network node over a transmission path between the nodes comprising: memory for, in use,
25 holding a database of activity distributions for particular classes of calls; means for classifying the call; means for selecting from the database a distribution appropriate to that class; means for determining a current load on the transmission path due to calls of that class; means for generating a posterior activity distribution from the selected distribution and current load; and means to reject or accept the
30 call on the basis of the posterior distribution.

4

A specific embodiment of the invention will now be described, by way of example only, with reference to the drawings in which:

5   Figure 1 shows a communication network, being controlled by a method in accordance with the invention;

Figure 2 is an explanatory diagram;

10   Figure 3 is a table data structure used in the communication network shown in Figure 1;

Figure 4 is a flow chart of the method used to control the network shown in Figure 1; and

15   Figure 5 is an explanatory diagram.

With reference to Figure 1, a communication network 1 comprises a plurality of nodes 2,3,4 and 5 interlinked by transmission paths 6, 7 and 8. The precise
20   nature of the nodes and transmission paths will depend on the network and they may be, for example, multiplexers, switches and optical fibre. The operation of the network is governed by a network manager 9 which may itself be a node (or distributed over a number of nodes) comprising a computer terminal of known type such as a Sun Microcomputer, marketed by the Sun Corporation which monitors,
25   communicates with, and controls the nodes 2, 3, 4 and 5 along control data paths 10, 11, 12 and 13 of known type. The network manager has a historical database 14 containing data about mean to peak bit rate distributions of call types having particular declared peak bit rate and a measured load datastore 9a. A telephone 5a is served by node 5 via a transmission path 5c.
30

Two typical mean to peak bit rate distributions are shown in Figure 2. The distributions can be thought of as a measure of the activity of a call in the sense of the likelihood of the call having a particular mean to peak bit rate. It will be seen

that distribution 31 is the distribution for a call class or type having a greater
activity than that of distribution 32. This is because its most probable mean to
peak bit rate, the peak of distribution 31, is closer to unity than that of distribution
32.

5

In practice, a large number of prior distributions will be held on the database 14 in
the form of a table 41 having a data-structure, as is illustrated in Figure 3. The
table 41 comprises two fields, a first data field 42 called PEAK and a second data
field 43 called DISTRIBUTION.

10

The PEAK field 42 contains ranges of possible peak bit rate values which will be
declared by a node when a request is made to place a call on the network.

DISTRIBUTION field 43 will contain parameterised versions of the distribution
15 curves appropriate to each range of peak bit rates stored in the PEAK field 42.

The network manager 9 operates to control the network 1 in a manner which will
now be described.

20 Consider the network 1 as it operates under the following conditions. Node 4 is
communicating with node 3 and node 2. The call traffic is a mixture of call types
such as constant bit rate and variable bit rate voice and video data. The network
manager periodically measures the load on the transmission path due to the call
classes, as represented by box 21 in the flow chart shown in figure 4. The
25 determined load is stored in datastore 9a, as represented by box 22. The network
manager 9 then monitors the network for a call connect request box 23. If a call
connect request is not received after a time delay, as represented by box 24, the
load is measured again as represented by box 21.

30 Node 5, not at present transmitting data to other nodes, wishes to do so and
requests the network manager 9 to allow connection to node 4. This is
represented by box 25 of the flow chart shown in Figure 4. When requesting
connection, node 5 declares its peak bit rate which is determined by the type of

equipment of which the node 5 is comprised and will thus be *"known"* to the node.

The declared peak is then compared with the value ranges stored in the historical
5   database 14, this in effect classifying the call, as represented by box 26, and an
appropriate historical mean to peak bit rate distribution is selected, as represented
by box 27.

The network manager 9, measures the network load as represented by box 28 but
10   in this case it is the load for the class of the call which wishes to be accepted. In
particular, it monitors the load on a transmission path between the transmitting
and receiving nodes, in this case path 6.

The measured load and the selected distribution are then used by the network
15   manager 9 to generate a posterior distribution, as represented by box 29. The
preferred way of generating the posterior distribution is by a Bayesian analysis and
it is a distribution of mean to peak bit rate against probability as earlier described,
but reflecting the predicted behaviour of the current call load for a particular class.

20   The network manager 9 then uses the posterior distribution in a call acceptance
algorithm, as represented by box 30, to determine whether the call is to be
accepted, as represented by box 31, or rejected, as represented by box 32.

The preferred method for determining whether or not to accept the call onto the
25   transmission path is by determining the effective bandwidth of the present call
load as will now be described.

The transmission path 6 has a capacity C and supports a number of calls separated
into I different classes. The calls in each class share the same characteristics. If
30   the effective bandwidth of a call in the j th class is aj and there are nj, such calls
then the quality of service is maintained only if:

$$C \geq \sum_{j=1}^{I} aj.nj$$

Thus, if the network manager 9 derives from the posterior distribution an estimate of the mean bit rate of the load on the transmission path 6. The estimated mean bit rate is then used with the peak bit rate for that class of call to calculate the effective bandwidth in the following manner.

If in the Bayesian analysis, the peak bit rate of the call is normalised to 1 and the prior distribution $\pi(p)$ represents the original belief of the activity (p) of a call defined as the mean bit rate to peak bit rate ratio.

If the load generated by the n calls (of this type) currently carried by the link is represented by S_n and in the measurement period is s.

If $P(S_n = s|p)$ is the probability of the load S_n being s in the last measurement period, conditional on the activity of an individual call being p.

The posterior distribution $\pi(p|S_n = s)$ for the activity is then given by,

$$\pi(p|S_n = s) = \frac{\pi(p)P(S_n = s|p)}{\int_0^1 \pi(p)P(S_n = s|p)dp}$$

The cell loss probability L is defined as,

$$L = \frac{Expected\ number\ of\ cells\ lost}{Expected\ number\ of\ cells\ sent}$$

5   The estimate for the cell loss probability $\hat{L}(n,s)$, from n calls, supported by a resource of capacity C is given by,

$$\hat{L}(n,s) = \frac{E\left[S_n(t) - C \mid S_n(0) = s\right]}{nE\left[X(t) \mid S_n(0) = s\right]}$$

10

Where $S_n(0)$ represents the monitored load used to obtain the posterior distribution, $S_n(t)$ the load from all calls of this type at time t and $X(t)$ the load produced by an individual call at time t.

15 This can be expressed as:

$$\hat{L}(n,s) = \frac{\int_0^1 \sum_{m=1}^{n-c} P\{S_n(t) \geq C + m \mid p\} \pi(p \mid S_n(0) = s) dp}{n \int_0^1 p\pi(p \mid S_n(0) = s) dp}$$

If the maximum permissible cell loss is L_{max}, the maximum number of calls of this

20 type which can be multiplexed n_{max} must therefore obey,

$$L_{max} \geq \hat{L}(n,s) = \frac{\int_0^1 \sum_{m=1}^{n_{max}-c} P\{S_{n\,max}(t) \geq C + m \mid p\} \pi(p \mid S_n(0) = s) dp}{n \int_0^1 p\pi(p \mid S_n(0) = s) dp}$$

5

This can be evaluated using the effective approximation,

$$10 \quad P\{S_{n\max}(t) \geq C \mid p\} \approx \exp\left\{-n_{\max}\left(\frac{c}{n_{\max}}l_n\frac{c}{n_{\max}p}+\left(1-\frac{c}{n_{\max}}\right)l_n\left(\frac{1-\frac{c}{n_{\max}}}{1-p}\right)\right)\right\}$$

for $n_{\max}P < C$ and

$$P\{S_{n\max}(t) \geq C \mid p\} = 1$$

for $n_{\max}P \geq C$ and

15

From $n_{n\max}$ the effective bandwidth α for the call is obtained by using,

$$\alpha = \frac{c}{n_{\max}}$$

20

This effective bandwidth is added to the effective bandwidth of the stored network load retrieved from database 9a by the network manager 9. The total effective bandwidth is then compared with the capacity of the transmission path 6 and if the required quality of service is maintained then the call is allowed onto the 25   transmission path 6.

If the required quality of service is not maintained, then an alternative transmission path is selected and considered as before.

In alternative embodiments, the load on the transmission path may be determined by providing a database storing tables of effective bandwidth for particular classes of calls for particular loads and numbers of calls. The network manager will then obtain the effective bandwidth by inputting the number of calls of the various
5   classes on the transmission path in question.

The provision of load values as a table structure will reduce the amount of processing time required since the values are not generated on-line. This will result in the need for a larger memory capacity than would be necessary if the values
10 were generated on-line.

In effect, an acceptance boundary is utilised which may be represented by a curve on a graph of element load against number of connections for a particular class of call as shown in figure 5. Curve A represents the situation where exact prior
15 information is known about the mean bit rates of equipment attempting to establish connections to the network. Curve B represents the situation where the information is only vague. It can be seen that where the exact prior information is known only eighty calls are accepted to a node. Where less information is known as the element load increases (goes to unity) only a small number of calls (forty)
20 are accepted. As the measure load decreases the algorithm allows more calls to be accepted since the calls generate less traffic.

Curve B can be held as a function or as a set of parameterised values in a look-up table. Thus a value of the current element load can be input into the table and an
25 allowed number of connections output from the table. If this number of connections is not exceeded then a further call can be accepted onto the network.

For example, if the current element load is Y then sixty connections can be allowed onto that element.
30

The prior distributions held in the historical database may, in some embodiments, be updated by the network manager from the monitored load.

If the call is rejected for a particular transmission path, the network manager may find an alternative route for the call in a manner well known to those skilled in the art. The new route will then be tested in the same way by generating a posterior distribution from the selected historical distribution and the new transmission
5 path's load.

CLAIMS

1.   A communication network control method for controlling acceptance
5   of a call from a first network node to a second network node over a transmission
path between the nodes comprising providing a database of activity distributions
for particular classes of calls, classifying the call, selecting from the database a
distribution appropriate to the call, determining a current load on the transmission
path due to calls of that class, generating a posterior activity distribution from the
10 selected distribution and current load, and rejecting or accepting the call on the
basis of the posterior distribution.

2.   A method as claimed in Claim 1 wherein the activity distributions are
mean to peak bit rate distributions.
15

3.   A method as claimed in Claim 1 or Claim 2 wherein the posterior
distribution is generated from the selected distribution and current load by a
Bayesian analysis.

20 4.   A method as claimed in Claim 1, Claim 2 or Claim 3 wherein the call
is classified by reference to a declared peak bit rate for the call.

5.   A method as claimed in any one of Claims 1, 2, 3 or 4 wherein an
effective bandwidth of calls is determined from the posterior distribution and the
25 effective bandwidth compared with the bandwidth of the transmission path to
determine whether to accept or reject the call for the transmission path.

6.   A method as claimed in Claim 5 wherein the effective bandwidth is
used to determine a predicted quality of service if the call is accepted onto the
30 transmission path and the predicted quality of service is compared with a required
quality of service and the call is rejected if the predicted quality of service is less
than the required quality of service.

7.   A method as claimed in any preceding claim wherein the load on the transmission path for all call classes is monitored and a measure of the monitored load stored and the call is accepted of the basis of stored monitored load and the posterior distribution.

8.   Apparatus for controlling acceptance of a call from a first network node to a second network node over a transmission path between the nodes which apparatus comprising: memory for, in use, holding a database of activity distributions for particular classes of calls; means for classifying the call; means for selecting from the database a distribution appropriate to that class; means for determining a current load on the transmission path due to calls of that class; means for generating a posterior activity distribution from the selected distribution and current load; and means to reject or accept the call on the basis of the posterior distribution.

9.   Apparatus as claimed in claim 8 wherein the activity distributions are mean to peak bit rate distributions.

10.   Apparatus as claimed in claim 8 or claim 9 wherein the means for generating a posterior activity distribution includes means to perform a Bayesian analysis.

11.   Apparatus as claimed in any one of claims 8, 9 or 10 wherein the means for classifying a call includes a look-up table of peak-bit rates and corresponding activity distributions.

12.   Apparatus as claimed in any preceding claim wherein the call is accepted or rejected by the means for accepting or rejecting a call determining an effective bandwidth from the posterior distribution and comparing the effective bandwidth with the bandwidth of the transmission path to determine whether to accept or reject the call.

13.   Apparatus as claimed in any preceding claim comprising means to, in use, store a required quality of service; means for predicting a quality of service if the call is accepted; and means for comparing the predicted quality of service with the required quality of service, the call being rejected if the predicted quality of

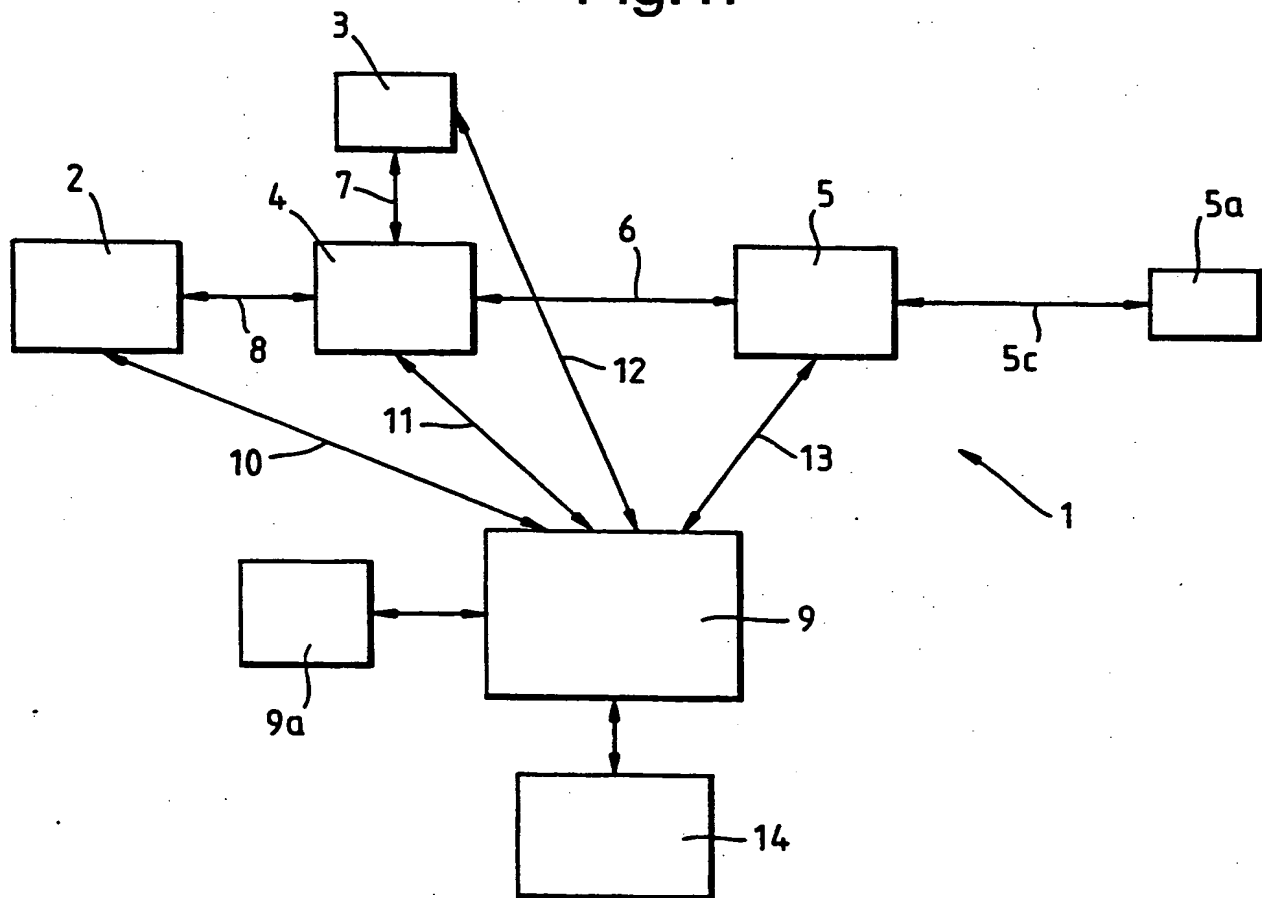5   service is less than the required quality of service.
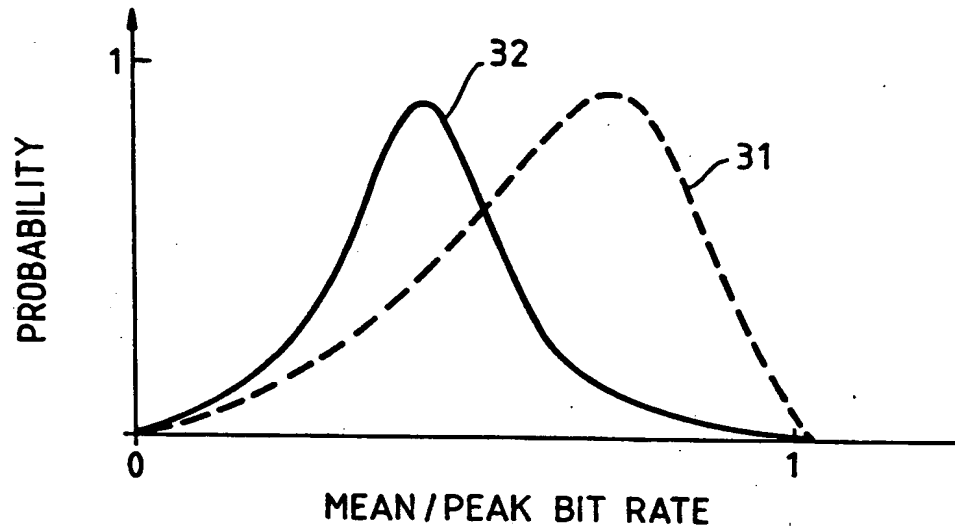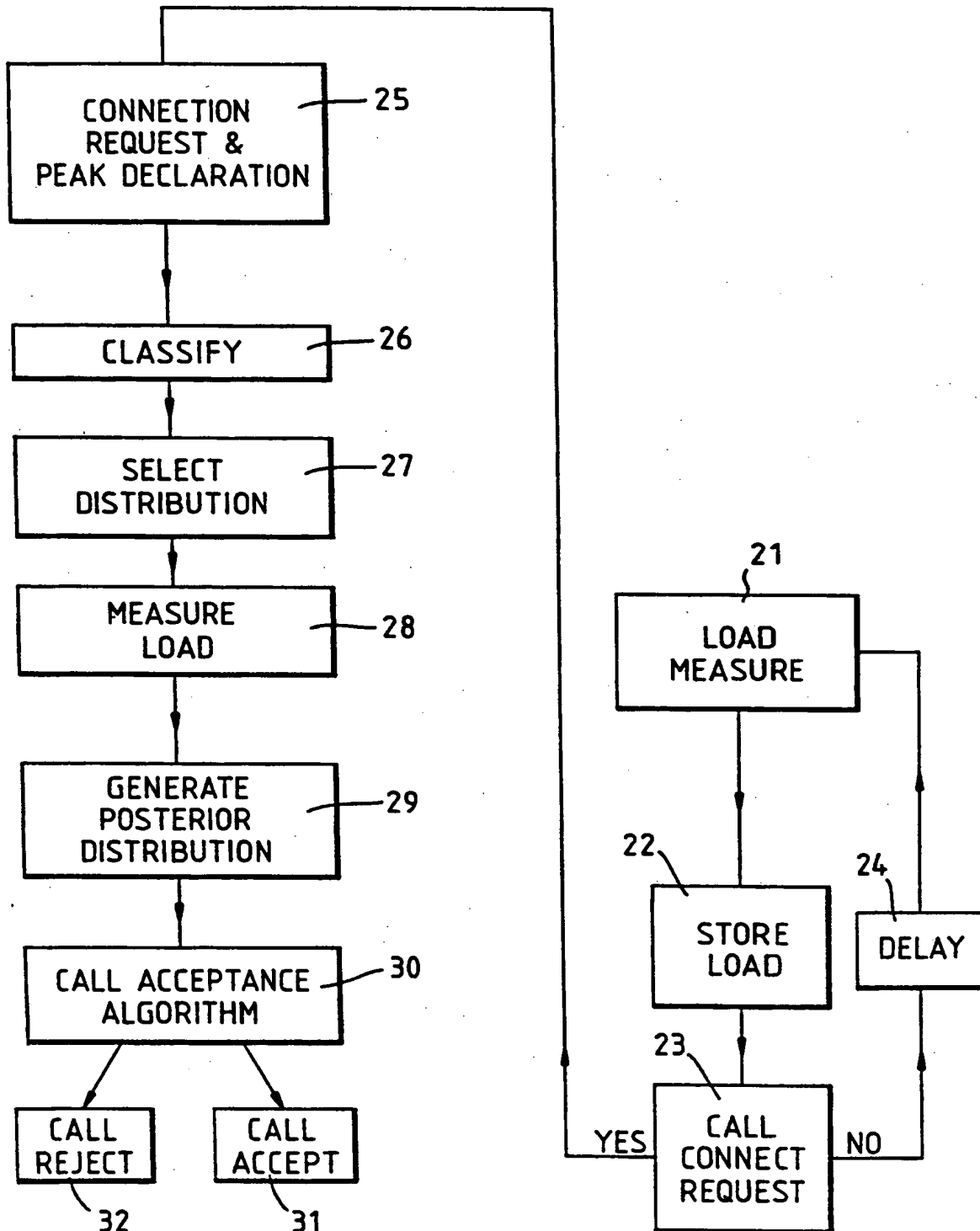
10

15

20

25

30

Fig.1.

Fig.2.

Fig.3.

Fig.4.

CONNECTION
REQUEST &
PEAK DECLARATION — 25

CLASSIFY — 26

SELECT
DISTRIBUTION — 27

MEASURE
LOAD — 28

GENERATE
POSTERIOR
DISTRIBUTION — 29

CALL ACCEPTANCE
ALGORITHM — 30

CALL
REJECT — 32

CALL
ACCEPT — 31

LOAD
MEASURE — 21

STORE
LOAD — 22

DELAY — 24

CALL
CONNECT
REQUEST — 23

YES   NO

Fig.5.

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   H04L12/56   H04Q11/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   H04L   H04Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GLOBECOM '91, vol. 1 OF 3, 2 December 1991 PHOENIX, pages 331-338, XP 000325994 R.BOLLA,F.DAVOLI,A.LOMBARDO,S.PALAZZO,D.PA NNO 'ADAPTIVE ACCESS CONTROL OF MULTIPLE TRAFFIC CLASSES IN ATM NETWORKS. 325994' see page 331, right column, line 40 - page 332, left column, line 25 ---	1-13
Y	EP-A-0 433 699 (SIEMENS) 26 June 1991 see column 11, line 30 - line 47; claim 1 --- -/--	1-13

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 July 1995	1 8. 07. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Veen, G

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	IEEE TRANSACTIONS ON INFORMATION THEORY, vol. 38, no. 1, January 1992 NEW YORK, pages 177-182, XP 000246277 Y.LIU,S.D.BLOSTEIN 'Optimality of the Sequential Probability Ratio Test for Nonstationary Observations' see page 178 - page 179	3,10
A	IEEE INFOCOM'92, vol. 1 OF 3, 4 May 1992 FLORENCE, pages 95-104, XP 000300050 P.SKELLY,S.DIXIT 'A Histogram-based Model for Video Traffic Behavior in an ATM Network Node with an Application to Congestion Control' see page 98, right column, line 37 - page 100, right column, line 7	1-13
A	IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, vol. 9, no. 7, September 1991 NEW YORK, pages 968-981, XP 000272772 R.GUERIN,H.AHMADI,M.NAGHSHINEH 'Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks.' see page 970, left column, line 1 - line 29	6,13

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0433699	26-06-91	NONE	